

AMENDMENTS TO THE SPECIFICATION

[0041] Figure 6 is a data structure diagram showing the data tables that preferably underlie the sample dimensional model shown in Figure 5. The diagram shows fact table 610, as well as dimension tables 620, 630, 640, and 650. The fact table 610 has columns 611-614, each corresponding to the key of one of the dimension tables. For example, sample row 616, corresponding to the sample company root page whose attribute values are shown in Figures 3 and 4, contains the company_name_key 237, as does row 626 of dimension table 620 for the company_name dimension. That row of dimension table 620 indicates that the company root page to which row 616 of the fact table corresponds has the value "Hughes Satellite, Inc." for the company_name attribute. The fact table 610 further includes column 615, which contains in each row the URL of the company root page to which the row corresponds, in this case, the root page of the Hughes Satellite company"http://www.hughessatellite.com". When the facility obtains the attribute information for this company root page, it adds row 616 to the fact table 610. It further searches each of the dimension tables to determine whether the obtained attribute value for that dimension is already contained in the dimension table. If so, the facility merely copies the key for that row of the dimension table into the row being added to the fact table. If the attribute value is not already in the dimension table, then the facility adds a row to the dimension table containing the attribute value and a key that is unique within the dimension table, and adds that key to the row being added to the fact table.